

Techniki umożliwiające komunikację mmWave dla systemów bezprzewodowych 5G: perspektywa na poziomie łącza

WPROWADZENIE

Gwałtowny wzrost usług multimedialnych wymagających dużej przepustowości i chęć powszechnego dostępu do danych przez użytkowników mobilnych zmusiły operatorów telekomunikacyjnych do ponownego przemyślenia sposobu projektowania sieci komórkowych. Aby sprostać wymaganiom, branża bezprzewodowa opracowała nowy standard piątej generacji (5G). Badania nad systemami bezprzewodowymi 5G nowej generacji, których celem jest rozwiązanie kilku bezprecedensowych wymagań technicznych i wyzwań, przyciągają w ciągu ostatnich kilku lat rosnące zainteresowanie zarówno środowiska akademickiego, jak i przemysłu. Aby spełnić wymagania 5G, można pomyśleć o poprawie wydajności widmowej i energetycznej obecnej sieci czwartej generacji (4G), która jest głównie przepelniona z 600 MHz do 3 GHz. Jednak badania pokazują, że sama taka poprawa nie zapewni przepustowości sieci 5G. Z tego powodu pasma częstotliwości mmWave w zakresie od 30 do 300 GHz wzbudziły duże zainteresowanie, jeśli chodzi o spełnienie wymagań przepustowości sieci 5G, gdzie dostępna jest ogromna ilość przepustowości. Chociaż dostępne pasmo częstotliwości mmWave jest bardzo duże, charakterystyka propagacji znacznie różni się od charakterystyki pasm częstotliwości mikrofalowych, które podsumowano w skrócie w następujący sposób:

1. Utrata ścieżki: z prawa Friisa izotropowa utrata ścieżki rośnie wraz z częstotliwością nośną. Na przykład, tłumienie ścieżki w wolnej przestrzeni maleje wraz z kwadratem częstotliwości nośnej. Zatem w komunikacji punkt-punkt można spodziewać się znacznej utraty ścieżki, gdy przejdziemy z częstotliwości nośnej 3 do 60 GHz.
2. Dyfrakcja i blokowanie: Dyfrakcja prowadzi do propagacji fal w geometrycznym obszarze cienia za przeszkodami. Dyfrakcja może spowodować niemożliwą do pominięcia propagację wielościeżkową zarówno w warunkach linii wzroku (LOS), jak i bez linii wzroku (NLOS). Z teorii elektromagnetycznej dobrze wiadomo, że fale elektromagnetyczne mają trudności z ugięciem, gdy propagują się na przeszkodach o wymiarach fizycznych znacznie większych niż długość fali [5]. Ponadto sygnały o częstotliwościach mikrofalowych mogą łatwiej przenikać przez ciała stałe i budynki niż mmWave. Z tych powodów na sygnały mmWave wpływa efekt cieniowania i dyfrakcji w znacznie większym stopniu niż sygnały mikrofalowe. Na przykład można zaobserwować straty zatorowe powyżej 35 dB spowodowane cegłami, betonem itp. Oraz około 35 dB spowodowane przez ciało ludzkie, gdzie straty te są pomijalne w paśmie częstotliwości mikrofalowych.
3. Tłumienie deszczu: Ogólnie straty spowodowane tłumieniem deszczu w pasmach częstotliwości mmFal są znacznie większe niż w pasmach mikrofalowych. Jeśli weźmiemy pod uwagę typową częstotliwość fal mm 73 GHz, można zaobserwować tłumienie deszczu o około 10 dB / km, które jest dość duże.
4. Absorpcja atmosferyczna: Wyniki pomiarów terenowych wykazały, że sygnały mmWave są bardziej podatne na pochłanianie tlenu niż sygnały mikrofalowe. Na przykład, można zaobserwować około 20 dB straty w okolicach 60 GHz sygnału mmWave.
5. Utrata liści: Tłumienie sygnałów radiowych spowodowane obecnością drzew blokujących łącze radiowe nazywane jest utratą liści [10]. Straty liści dla fal mmWaves są znaczące i mogą być czynnikiem ograniczającym w niektórych środowiskach propagacji. Wyniki empiryczne pokazują, że przy penetracji listowia na 10 m strata przy częstotliwości 80 GHz mm wynosi około 23,5 dB, czyli o około 15 dB więcej w porównaniu z częstotliwością mikrofalową 3 GHz.

Wszystkie te dowody potwierdzają, że całkowite straty w systemach mmWave są większe niż w systemach mikrofalowych. Na szczęście małe długości fal sygnałów mmWave umożliwiają umieszczenie dużej liczby elementów anteny w tym samym fizycznym obszarze anteny, zapewniając w ten sposób duże zyski przetwarzania przestrzennego, które teoretycznie mogą kompensować co najmniej odpowiednią izotropową utratę ścieżki. Jednakże, ponieważ systemy mmWave są wyposażone w kilka anten, pojawia się szereg wyzwań związanych z obliczeniami i implementacją, aby utrzymać oczekiwany wzrost wydajności. W tym celu w tym rozdziale omówiono kluczowe techniki umożliwiające komunikację mmWave dla sieci 5G z perspektywy poziomego łącza. Wydajność na poziomie łącza systemu bezprzewodowego mmWave zależy od wielu podejść umożliwiających, w tym schematu transmisji (tj. Czy stosujemy formowanie wiązki, multipleksowanie, czy oba), podejścia do identyfikacji kanału, sposobu projektowania struktury fali przesyłanego sygnału i strategii dostępu, które zostały pokrótce podsumowane w następujący sposób. Wiele wejść i wyjść wielokrotnych (MIMO) jest jedną z obiecujących technik poprawy wydajności kanałów bezprzewodowych na poziomie łącza. Aby w pełni wykorzystać potencjał MIMO, można wykorzystać konwencjonalne cyfrowe formowanie wiązki (DB), w którym do każdej anteny dedykowany jest jeden łańcuch częstotliwości radiowych (RF). W rzeczywistości DB jest realistycznym podejściem projektowym dla poprzednich i obecnych systemów bezprzewodowych, ponieważ systemy te są wyposażone w kilka anten (około 10) [11, 12]. Ponieważ jednak systemy mmWave są prawdopodobnie wyposażone w duże tablice anten, przeznaczenie oddzielnych łańcuchów RF dla każdej anteny jest niewykonalne. To motywuje naukowców do stworzenia systemu mmWave, w którym liczba łańcuchów RF jest znacznie mniejsza niż w antenach. Jednym ze sposobów realizacji takiego systemu jest wdrożenie hybrydowego formowania wiązki (HB), gdzie jest realizowane zarówno w domenie analogowej, jak i cyfrowej (tj. Hybrydowe formowanie wiązki analogowo-cyfrowe). Komunikacja MIMO jest również dobrze znana ze swojej zdolności do przesyłania niezależnych informacji przez różne elementy anteny (tj. Multipleksowanie przestrzenne). W szczególności systemy mmWave są odpowiednie do rozmieszczenia dużej liczby anten (rzędu 100–1000). Ponadto działają w bardzo wysokich pasmach częstotliwości, w których dostępne szerokości pasma są znacznie duże (około 7 GHz) przy 60 GHz. Zachęca to do rozważenia transmisji multipleksowanych przestrzennie dla kanałów mmWave w celu znacznego zwiększenia wydajności na poziomie łącza. Multipleksowanie przestrzenne zapewnia wzmocnienie multipleksowania, które zwiększa przepustowość transmisji poprzez podzielenie wychodzącego strumienia sygnału na wiele części, przy czym każdy element jest transmitowany jednocześnie i równoległe na tym samym kanale RF przez różne anteny. Ponadto, aby przestrzenna multipleksowana transmisja działała dobrze, kanał musi zapewniać wystarczającą dekorrelację między różnymi blisko rozmieszczonymi antenami. Multipleksowanie przestrzenne może być włączone dla systemu pojedynczego użytkownika poprzez wykorzystanie przestrzennej separacji anteny, a dla systemu wielu użytkowników poprzez wykorzystanie informacji kierunkowej wyposażenia użytkownika (UE). Wydajność formowania wiązki i multipleksowania przestrzennego zależy od dostępności informacji o stanie kanału (CSI), gdzie uzyskuje się lepszą pojemność, gdy akwizycja CSI jest przeprowadzana wydajnie. Ogólnie, jakość estymatora kanału można poprawić, zwiększając okres uczenia kanału MIMO. Jednakże, ponieważ kanały praktyczne mają ograniczony czas koherencji, gdzie kanał jest traktowany jako prawie stały, okres szkolenia nie może zostać wybrany zgodnie z życzeniem. Ponadto, ponieważ systemy mmWave są prawdopodobnie wyposażone w mniejszą liczbę łańcuchów RF niż anteny, może być konieczne zidentyfikowanie CSI ze zmniejszonymi wymiarami przestrzennymi, co w konsekwencji pogarsza jakość oszacowania kanału i wydajność poziomego łącza [22]. Kształt fali przesyłanego sygnału jest kolejnym ważnym czynnikiem wpływającym na wydajność poziomego łącza w kanale bezprzewodowym. W rzeczywistości różne projekty przebiegów prowadzą do różnych wydajności systemu pod względem estymacji kanału i formowania wiązki. Na przykład, jeśli wykorzystuje się dobrze znany kształt fali multipleksowania z ortogonalnym podziałem częstotliwości

(OFDM), taka konstrukcja wymaga wzmacniacza mocy (PA) o dużym zakresie dynamicznym dla każdego elementu anteny. Spowoduje to kolejny wzrost ceny dla systemów mmWave. Wręcz przeciwnie, jeśli używa się tanich PA, niektóre z transmitowanych sygnałów zostaną obcięte, co z kolei spowoduje utratę ortogonalności każdej podnośnej. Ponadto konstrukcja ta zapewnia bardzo wysoki stosunek mocy szczytowej do średniej (PAPR) dla przesyłanych sygnałów, co poważnie obniża wydajność systemu pod względem współczynnika błędów bitowych (BER) [25]. Innym ważnym czynnikiem wpływającym na wydajność na poziomie łącza systemów mmWave jest strategia dostępu do kanału bezprzewodowego. W związku z tym idea jednoczesnego podłączenia UE zarówno do pasm mikrofalowych, jak i pasm mmWave była promowana w przypadku nakładanych układów mmWave z małymi komórkami (tj. Mała komórka jest nakładana na stację bazową makro (BS)). W ten sposób płaszczyzny sterowania i danych mogą zostać rozdzielone, tak aby krytyczne dane sterujące były przesyłane przez niezawodne łącza mikrofalowe między UE i makro BS, podczas gdy szybka transmisja danych między UE a małokomórkowymi stacjami BS odbywa się w pasmach częstotliwości mmWave. W konsekwencji niezawodna i stabilna komunikacja może być utrzymana, czerpiąc korzyści z pasm mmWave. W kolejnych sekcjach szczegółowo omówimy wyżej wymienione techniki wspomagające dla przyszłej sieci 5G, z naciskiem na systemy mmWave i wynikającą z nich wydajność na poziomie łącza.

KSZTAŁTOWANIE PROMIENI

Rozmieszczenie wielu anten w nadajniku i / lub odbiorniku poprawia ogólną wydajność systemu. Ta poprawa wydajności jest możliwa dzięki wykorzystaniu formowania wiązki, gdzie jest to realizowane przez wykorzystanie CSI. Wydajność formatora wiązki osiąga maksymalne korzyści, gdy współczynniki kanału odpowiadające różnym parom anten nadawczo-odbiorczych ulegają niezależnemu zanikaniu. Dla danej częstotliwości nośnej taki niezależny zanikający kanał występuje, gdy odległość między dwoma elementami anteny wynosi co najmniej $0,5\lambda$, gdzie λ jest długością fali. Tak więc, dla stałego wymiaru przestrzennego, liczba rozmieszczonych anten rośnie wraz ze wzrostem częstotliwości nośnej, co w konsekwencji pomaga upakować duże tablice antenowe przy częstotliwościach mmWave w porównaniu z mikrofalami. Przykładowo, przy rozmiarze anteny 1m można zmieścić około 214 i 14 elementów antenowych, pracując na częstotliwości nośnej 30 i 2 GHz odpowiednio.

CYFROWE KSZTAŁTOWANIE WIĄZKI

Kształtowanie wiązki to proces kontrolowania amplitudy i fazy nadawanego i / lub odbieranego sygnału zgodnie z żądaną aplikacją i środowiskiem kanału. W systemie MIMO transmisja formowana wiązką jest najpowszechniej stosowanym podejściem, w którym można ją zrealizować zarówno w systemach jednego użytkownika, jak i wielu użytkowników. Szereg podejść do kształtowania wiązki jest dostępnych w kilku pracach badawczych i standardowych książkach komunikacyjnych MIMO. W konwencjonalnej DB, jeden łańcuch RF jest wymagany dla każdego elementu anteny w BS i UE, gdzie łańcuch RF zawiera wzmacniacz o niskim poziomie szumów, konwerter w dół, konwerter cyfrowo-analogowy (DAC), konwerter analogowo-cyfrowy (ADC) i tak dalej. Aby pokazać podstawy DB, zbadamy scenariusz, w którym stacja BS z N anten obsługuje K zdecentralizowanych UE z pojedynczą anteną z parametrami podsumowanymi w tabeli 9.2.

Zmienne: Definicja

d_k : przesłany symbol dla k-tego UE na okres symbolu

$d = [d_1, d_2, \dots, d_k]$: Przesłany symbol wszystkich UE na okres symbolu

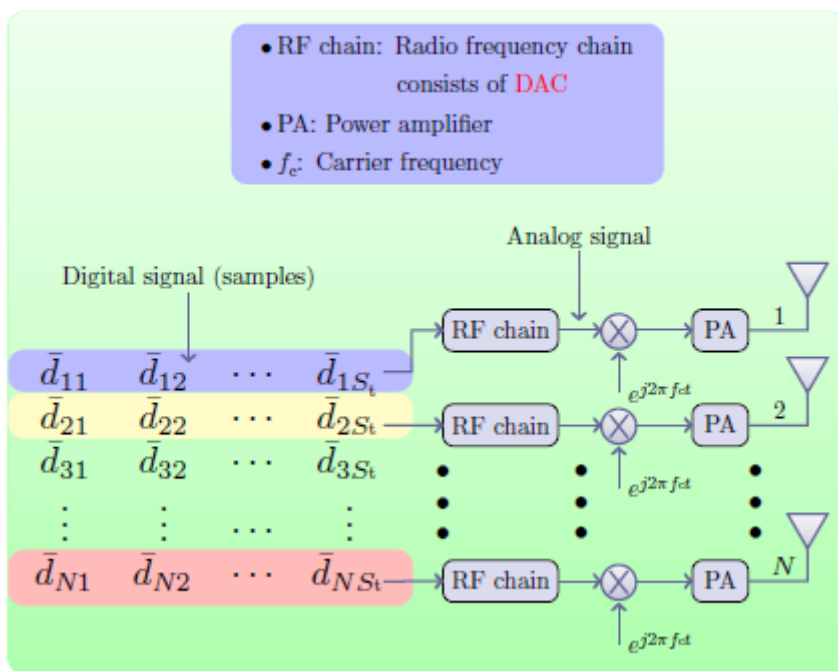
$b_k \in \mathbb{C}^{N \times 1}$: wektor kształtujący wiązkę dla k-tego UE

$B = [b_1, b_2, \dots, b_k]$: Macierz kształtowania wiązki dla wszystkich UE

$x_k = b_k d_k$: próbka kształtowana wiązką dla k-tego UE na okres symbolu

$\bar{\mathbf{d}} = \sum_{i=1}^K \mathbf{x}_i = \mathbf{B}\mathbf{d}$: Przesyłane próbki na okres symbolu

W każdym czasie trwania symbolu, symbole danych wszystkich UE są najpierw łączone w postaci wektorowej $\mathbf{d} = [d_1, d_2, \dots, d_K]$, gdzie d_k jest symbolem danych k-tego UE. Następnie k-ty symbol UE zostanie pomnożony przez odpowiadający mu wektor kształtujący wiązkę b_k , aby utworzyć przesyłany wektor odpowiadający k-temu UE jako $x_k = b_k d_k$. Na koniec wszystkie transmitowane wektory zostaną zsumowane jako $\bar{\mathbf{d}} = \sum_{i=1}^K \mathbf{x}_i = \mathbf{B}\mathbf{d}$, przepuszczone przez łańcuchy RF i przesłane do powietrza przez wszystkie anteny. Teraz, jeśli chcielibyśmy transmitować symbole S_t dla każdego UE, nadajnik nadaje sekwencje informacji $\bar{\mathbf{d}}_1, \bar{\mathbf{d}}_2, \dots, \bar{\mathbf{d}}_{S_t}$. Proces ten obrazowo przedstawiono na rysunku

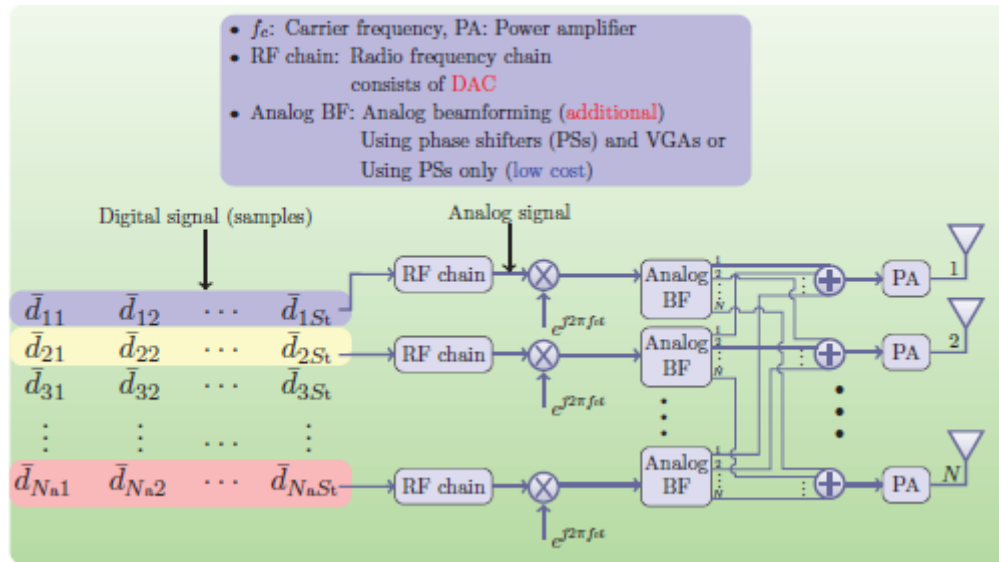


Ogólnie, macierz kształtowania wiązki \mathbf{B} jest parametrem projektowym zależnym od kanału i szumu, który może wymagać optymalizacji w oparciu o różne kryteria projektowe. Jak widać na tym rysunku, DB wymaga, aby każda antena była podłączona do jednego łańcucha RF. DB jest powszechnie stosowany do zastosowań częstotliwości mikrofalowych, ponieważ liczba rozmieszczonych anten jest rzędu 10 (np. Obecna sieć LTE). Jednakże, ponieważ systemy mmWave są prawdopodobnie wyposażone w kilka anten, a koszt i pobór mocy ADC i DAC jest bardzo wysoki, poświęcenie oddzielnego łańcucha RF dla każdej anteny jest praktycznie niewykonalny. Z tego powodu projekt kształtowania wiązki z ograniczoną liczbą łańcuchów RF zyskał ostatnio dużą uwagę.

HYBRYDOWE TWORZENIE WIĄZKI

Poświęcenie oddzielnego łańcucha RF dla każdej anteny nie jest ekonomiczne w przypadku zastosowań z dużymi antenami, takich jak mmWave. To skłania naukowców do myślenia o systemie, w którym liczba łańcuchów RF jest znacznie mniejsza niż liczba anten. Jednym ze sposobów realizacji takiego systemu jest wdrożenie HB przy użyciu zarówno domeny analogowej, jak i cyfrowej. W domenie cyfrowej formowanie wiązki może być realizowane za pomocą mikroprocesorów (superkomputerów),

podczas gdy w domenie analogowej formowanie wiązki jest realizowane przy użyciu przesuwników fazowych (PS) i ewentualnie wzmacniaczy o zmiennym wzmacnieniu (VGA). Rysunek poniższy przedstawia typową architekturę nadajnika dla systemu, w którym stacja BS mająca N anten i łańcuchy $N_a < N$ RF zawierająca hybrydowe analogowo-cyfrowe formowanie wiązki obsługuje K zdecentralizowanej pojedynczej anteny UE, a wszystkie inne ustawienia są takie, jak pokazano w tabeli



W tej hybrydowej architekturze podsumowano operację formowania wiązki w tabeli 9.3.

Zmienne: Definicja

d_k : przesłany symbol dla k-tego UE na okres symbolu

$\mathbf{d} = [d_1, d_2, \dots, d_k]$: Przesłany symbol wszystkich UE na okres symbolu

$\tilde{\mathbf{b}}_k \in \mathbb{C}^{N_a \times 1}$: wektor kształtowania wiązki dla k-tego UE ($N_a < N$)

$\tilde{\mathbf{B}} = [\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2, \dots, \tilde{\mathbf{b}}_k]$: Macierz kształtowania wiązki dla wszystkich UE

$\tilde{\mathbf{x}}_k = \tilde{\mathbf{b}}_k d_k$: Próbką kształtowaną wiązką dla k-tego UE na okres symbolu

$\mathbf{A} \in \mathbb{C}^{N \times N_a}$: Analogowa macierz kształtowania wiązki wszystkich UE

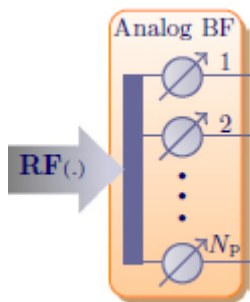
$\tilde{\mathbf{d}} = \mathbf{A} \sum_{i=1}^K \tilde{\mathbf{x}}_i = \mathbf{A} \tilde{\mathbf{B}} \mathbf{d}$: Przesyłane próbki na okres symbolu (A jest wspólne dla wszystkich UE)

Ponieważ cena łańcuchów RF jest znacznie wyższa niż cena analogowych komponentów elektronicznych przy realizacji analogowej macierzy formującej wiązkę A w HB, można kontrolować koszt projektu HB, kontrolując liczbę łańcuchów RF N_a . Ponieważ jednak hybrydowa architektura analogowo-cyfrowa straci stopni swobody $N - N_a$ w porównaniu z architekturą cyfrową, wydajność HB nie może być lepsza niż DB. W związku z tym kompromisem między hybrydowym i cyfrowym kształtowaniem wiązki jest koszt w porównaniu z wydajnością. Hybrydową architekturę analogowo-cyfrową po stronie odbiornika można również zaprojektować podobnie jak na rysunku powyżej, a jej szczegóły zostały pominięte dla zwięzłości. Część architektury HB, kształtującą wiązkę analogową,

można zrealizować przez wykorzystanie VGA i PS. Ponieważ jednak cena VGA jest znacznie wyższa niż PS, to projekt analogowego kształtowania wiązki przy użyciu tylko PS wzbudza duże zainteresowanie. Z tych powodów w ciągu ostatnich kilku lat zaproponowano kilka architektur HB, zakładając, że część DB jest realizowana przy użyciu mikroprocesorów, podczas gdy część analogowa formująca wiązkę jest realizowana tylko przy użyciu PS. W tym celu jednym z fundamentalnych problemów hybrydowej architektury analogowo-cyfrowej jest zastosowanie jak najmniejszej liczby łańcuchów RF przy zachowaniu prawie takiej samej wydajności systemu, jak w przypadku technologii cyfrowej. Główne wyzwanie wynika z faktu, że $N_a < N$ i elementy analogowej macierzy formującej wiązkę A są projektowane przy użyciu samych PS o stałym module.¹ W tym zakresie podjęto szereg prac badawczych w celu wspólnego zaprojektowania macierzy formujących wiązki \tilde{B} i A dla kilku kryteriów projektowych i konfiguracji systemu. Dla ustalonego A można zaprojektować \tilde{B} , stosując techniki stosowane w konwencjonalnej DB. Pozostaje zatem pytanie, jak zaprojektować A . Ogólnie rzecz biorąc, naukowcy zasugerowali dwa podejścia projektowe do zaprojektowania A : pojedynczy PS i sparowany PS.

Podejście z pojedynczym PS

Pierwsze podejście próbuje zaprojektować A , ograniczając każdy z jego elementów tak, aby miał stałą amplitudę, jak pokazano na rysunku poniżej (tj. Podejście projektowe z pojedynczym PS).



Jak widać na tym rysunku, sygnał pochodzący z każdego łańcucha RF przejdzie przez N_p PS. Gdy $N_p = N$, sygnał transmitowany z każdego łańcucha RF będzie trafiał do każdej anteny. Z tego powodu konstrukcja HB jest określana jako w pełni połączona hybrydowa architektura analogowo-cyfrowa z pojedynczym PS. A gdy $N_p < N$, przesyłany sygnał każdego łańcucha RF będzie uczestniczył tylko w niektórych antenach i dlatego jest określany jako częściowo połączona architektura HB. Przy takim podejściu prawie wszystkie algorytmy HB wykorzystują metodę książki kodowej do optymalizacji A , gdzie jej główną ideą jest wybranie kolumn A ze zbioru predefiniowanych wektorów. Pod tym względem technika dopasowywania pościgu (MP) została powszechnie przyjęta w kilku pracach literaturowych. Ogólnie algorytm HB oparty na MP osiąga bardzo dobrą wydajność zarówno dla pojedynczego użytkownika, jak i dla wielu użytkowników ogromne systemy MIMO, gdy kanał ma bardzo mało rozpraszaczy. Jednak podejście HB oparte na MP nie jest w stanie osiągnąć takiej samej wydajności jak w przypadku BD dla ogólnych macierzy kanałów. W szczególności różnica w wydajności między HB i DB jest bardzo duża, gdy liczba rozpraszaczy jest znacznie większa niż liczba łańcuchów RF (zostanie to uzasadnione w następnym podrozdziale). Również w Liang i in. rozważany jest schemat wstępnego kodowania hybrydowego pojedynczego PS dla masowych systemów MIMO dla wielu użytkowników. W artykule zastosowano hybrydowe kodowanie wstępne wymuszające zero (ZF), które ma na celu zmaksymalizowanie sumarycznej szybkości wszystkich użytkowników. Żadne z takich podejść HB jest w stanie utrzymać taką samą wydajność jak DB dla innych konstrukcji HB z pojedynczym PS). W żadnej z tych prac nie jest jasne, jak można analitycznie określić ilościowo związek między

wynikami HB i DB. Potencjalną tego przyczyną jest to, że takie podejście będzie modyfikować fazę każdego elementu tylko A. Motywuje to naukowców do projektowania analogowej macierzy formującej wiązkę w projekcie HB tylko z PS, zapewniając jednocześnie możliwość kontrolowania zarówno amplitudy, jak i fazy A, co prowadzi do sparowanego podejścia do projektowania PS.

Sparowane podejście PS

Rozumiemy, że podejście do projektowania pojedynczego PS pomaga modyfikować fazę każdego elementu A. W konsekwencji jego wydajność jest degradowana i takie podejście nie jest w stanie analitycznie ilościowo określić związku między wydajnością HB i DB. W przypadku sparowanego PS, ponieważ elementy A mogą mieć różną amplitudę i fazę, celem staje się uzyskanie pewnego związku między projektami HB i DB. Można temu zaradzić, określając ilościowo liczbę łańcuchów RF i sparowanych PS, tak aby projekty HB i DB osiągały taką samą wydajność. Jednym ze sposobów określenia tych liczb jest rozważenie określonych kryteriów projektowych i modelu kanału (np. maksymalizacja szybkości sumy z modelem kanału z jednolitą macierzą liniową (ULA)). Jednak takie podejście projektowe nie jest elastyczne, ponieważ nie może być rozszerzone na inne kryteria projektowe i modele kanałów (np. dla maksymalizacji sumy ważonej maksymalizacji szybkości, problemów z maksymalnym i minimalnym sygnałem do zakłóceń plus szumem (SINR) i tak dalej). To motywuje badaczy do zajęcia się powyższym problemem poprzez rozważenie arbitralnych kryteriów projektowych. W rzeczywistości wydajność HB będzie taka sama jak DB, gdy nadawane i odbierane sygnały w obu przypadkach są takie same. W związku z tym może być konieczne wybranie optymalnego $\tilde{A}\tilde{B}$ jako:

$$\tilde{A}\tilde{B} = B = UQ \quad (1)$$

gdzie B i $\{A, \tilde{B}\}$ są zdefiniowane odpowiednio w tabelach powyższych, a $UQ = QR(B)$ jest rozkładem QR B z $U^H U = I$. Korzystając z teorii macierzy, można udowodnić, że równości w równaniu (1) nie może być spełnione, gdy $N_a \leq \text{rank}(B) \leq K$ dla dowolnego B, co uzasadnia, że $N_a \geq K$. Teraz, jeśli mamy $N_a = K$, można pomyśleć o ustawianiu bezpośrednio $A = U$ i $\tilde{B} = Q$. Ponieważ jednak amplituda każdego z elementów U niekoniecznie jest taka sama, nie jest jasne, w jaki sposób można zrealizować tę macierz używając tylko PS. Dlatego taka bezpośrednia wtyczka nie pomoże w projektowaniu architektury HB. Z tego powodu Zhang i inni po raz pierwszy wymyślił nową i sprytną metodę reprezentacji „dowolnego” wektora $v \in \mathbb{C}^{N \times 1}$ jako :

$$v = Wz \quad (2)$$

gdzie $W \in \mathbb{C}^{N \times 2}$ i $z \in \mathbb{C}^{2 \times 1}$ z $|W_{(i,j)}|^2 = 1, \forall i, j$, co prowadzi Zhanga, aby stwierdzić, że wydajność dowolnego DB można osiągnąć z HB, jeśli liczba łańcuchów RF jest co najmniej dwa razy większa od liczby strumieni danych (tj. Dwa razy większa od rzędu B) Jeśli zastosujemy tę technikę w architekturze HB, potrzeba maksymalnie łańcuchów RF $N_a = 2K$ i PS $2KN$, aby osiągnąć taką samą wydajność, jak w przypadku projektu DB. Jest to pokazane, W z równaniu (2) nie jest wyjątkowe. To skłania do dalszej redukcji liczby łańcuchów RF i PS poprzez wykorzystanie stopnia swobody w W. Główna idea Bogalea miała wykorzystać następującą zależność: dla dowolnej liczby rzeczywistej x z $-2 \leq x \leq 2$ jest to pokazane tak ,że:

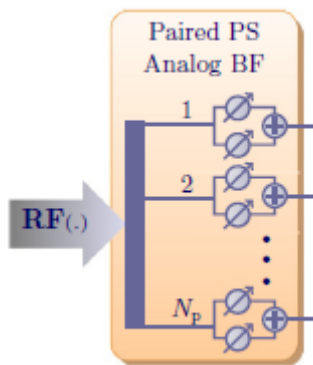
$$x = e^{j\cos^{-1}(\frac{x}{2})} + e^{-j\cos^{-1}(\frac{x}{2})}$$

$$jx = e^{j\sin^{-1}(\frac{x}{2})} + e^{j(\pi - \sin^{-1}(\frac{x}{2}))} \quad (3,4)$$

gdzie $j = \sqrt{-1}$. Równanie (3), można wyrazić (m, n) -ty element U jako:

$$\begin{aligned} U_{(m,n)} &\triangleq a_{mn} e^{j\phi_{mn}} \\ &= e^{j(\cos^{-1}(\frac{a_{mn}}{2}) + \phi_{mn})} + e^{-j(\cos^{-1}(\frac{a_{mn}}{2}) - \phi_{mn})} \end{aligned} \quad (5)$$

W konsekwencji każdy element U może być równoważnie wyrażony jako suma dwóch PS. Co więcej, ponieważ maksymalna liczba niezerowych elementów U to KN, rozwiązanie otrzymane w DB można osiągnąć stosując 2NK PS i łańcuchy Na K RF (tj. Zmniejszone o połowę w porównaniu z pracą Zhanga), co prowadzi do sparowanej architektury PS, jak pokazano na rysunku



Liczbę PS można nieznacznie zmniejszyć, a różne dostosowane wersje sparowanych PS można znaleźć u Bogalea. Ponadto taka sparowana architektura PS może być również w pełni połączona lub częściowo połączona.

WYDAJNOŚĆ NA POZIOMIE ŁĄCZA

W tej podsekcji przedstawiono wyniki symulacji w celu oceny wydajności na poziomie łącza DB i HB z projektami z pojedynczą architekturą PS. Do symulacji rozważamy system wielodostępny, w którym stacja BS jest wyposażona w anteny N ULA i K UE, w których każde UE jest wyposażone w anteny M_k ULA. Nadajnik i numer odbiornika dla łańcuchów RF jest ustawiony odpowiednio na P_t i P_k . Ponadto liczba rozpraszaczy między BS a k-tym UE jest ustawiona na L_k . Przy takich ustawieniach kanał między k-tym UE i BS można wyrazić jako :

$$\mathbf{H}_k = \sqrt{\frac{NM_k}{L_k \rho_k}} \sum_{i=1}^{L_k} g_{ki} \mathbf{a}_{tk}(\theta_{tk}(i)) \mathbf{a}_{rk}^H(\theta_{rk}(i)) \quad (6)$$

gdzie g_{ki} jest wzmocnieniem zespolonym k-tego UE i-tej ścieżki z $E\{|g_{ki}|^2\} = 1$, ρ_k jest tłumieniem ścieżki między nadajnikiem a k-tym odbiornikiem, $\theta_{t}(i) \in [0, 2\pi]$, $\theta_{r}(i) \in [0, 2\pi]$, $\forall i$ i $\mathbf{a}_t(\cdot)$ i $\mathbf{a}_r(\cdot)$ są wektorami odpowiedzi układu antenowego odpowiednio w nadajniku i odbiorniku. Tutaj przyjmujemy elementy anteny ULA, gdzie $\mathbf{a}_{tk}(\cdot)$ i $\mathbf{a}_{rk}(\cdot)$ są modelowane jako :

$$\mathbf{a}_{tk}(\theta) = \frac{1}{\sqrt{N}} [1, \exp^{j\frac{2\pi}{\lambda}d\sin(\theta)}, \dots, \exp^{j(N-1)\frac{2\pi}{\lambda}d\sin(\theta)}]^T$$

$$\mathbf{a}_{rk}(\theta) = \frac{1}{\sqrt{M_k}} [1, \exp^{j\frac{2\pi}{\lambda}d\sin(\theta)}, \dots, \exp^{j(M_k-1)\frac{2\pi}{\lambda}d\sin(\theta)}]^T$$

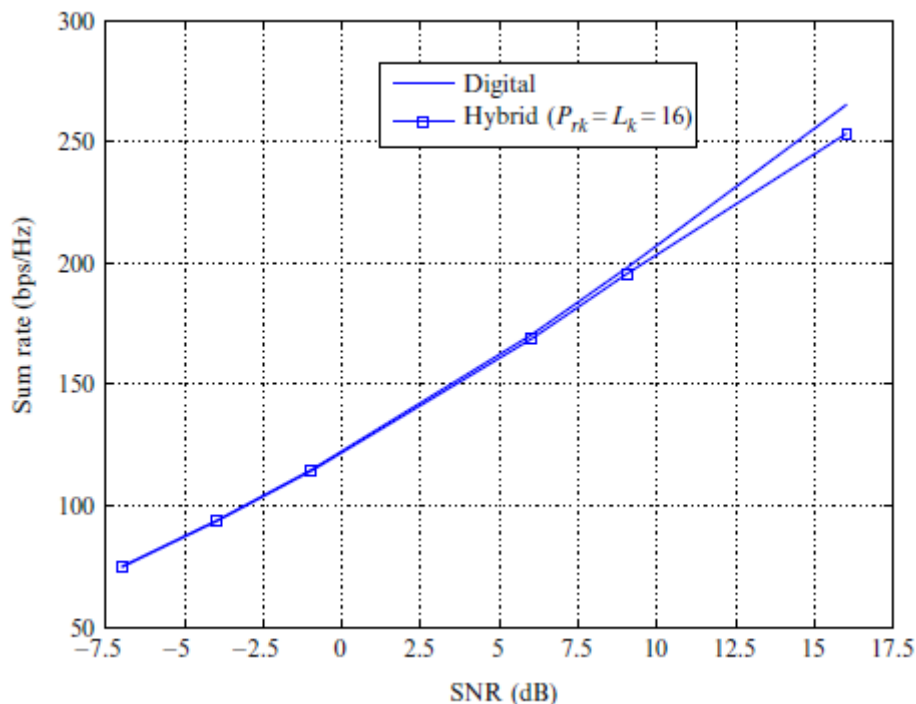
gdzie $j=\sqrt{-1}$, λ to długość fali nadawczej, a d to odległość między antenami. Użyliśmy $N = 128$, $M_k = 32$, $d = 0,5\lambda$, $K = 4$, $L_k = 16$ i $P_t = K P_{rk}$. SNR, który jest zdefiniowany jako $\text{SNR} = P_{av} / \sigma^2$ jest kontrolowane przez zmianę σ^2 przy zachowaniu całkowitej transmitowanej mocy $P_{\max} = K S_k m W$ i $P_{av} = P_{\max} / S_k$. Parametry kanału to $\rho = [0,2338 \ 0,2333 \ 0,0402 \ 0,5290]$, a $\theta_{tk}(i)$ ($\theta_{rk}(i)$) są pobierane losowo z jednorodnych zmiennych losowych w $[0, 2\pi]$ (patrz równanie 9.7). Całkowitą sumaryczną stawkę oblicza się jako

$$R_t = \sum_{k=1}^K \sum_{i=1}^{S_k} R_{ki},$$

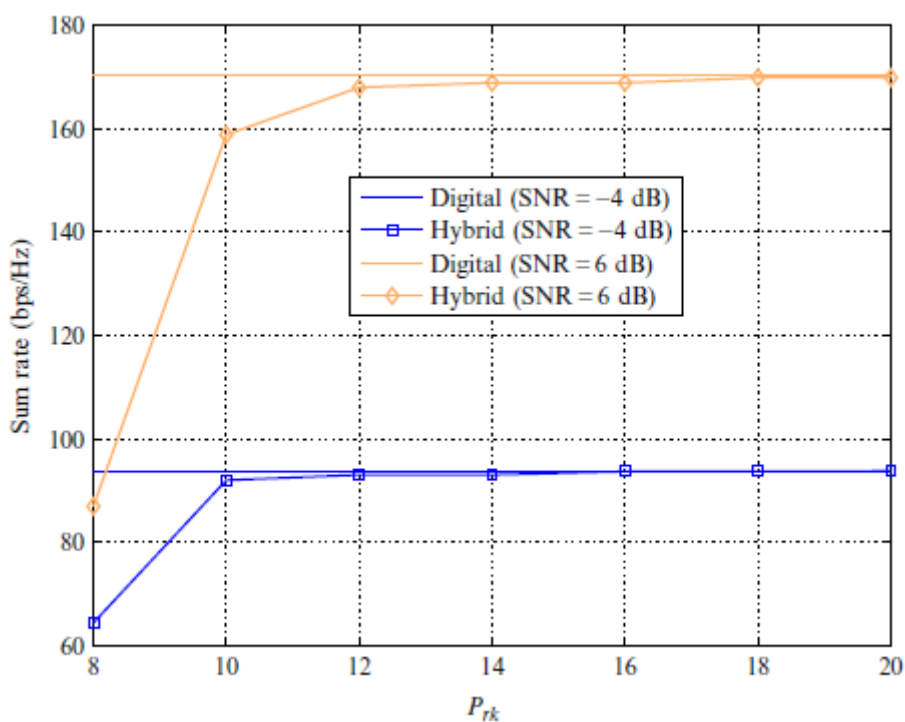
gdzie stawka każdego symbolu jest obliczana jako $R_{ki} = \log_2(1 + \gamma_{ki})$ z γ_{ki} jako osiągniętym SINR k-tego symbolu UE i-tego. Wszystkie wykresy są generowane przez uśrednienie 1000 realizacji γ_{ki} , $\forall k$:

$$\theta_t(\text{in } \pi) = \begin{bmatrix} 0.71 & 0.66 & 0.20 & 0.70 \\ 1.29 & 0.16 & 0.81 & 1.88 \\ 0.65 & 0.91 & 0.22 & 0.52 \\ 1.70 & 0.64 & 0.94 & 1.61 \\ 0.80 & 1.85 & 1.42 & 1.33 \\ 0.43 & 1.11 & 1.99 & 1.38 \\ 1.24 & 0.61 & 1.53 & 1.95 \\ 0.80 & 0.10 & 0.77 & 1.22 \\ 0.47 & 0.07 & 0.38 & 0.39 \\ 0.43 & 1.50 & 0.83 & 1.83 \\ 0.04 & 0.71 & 1.27 & 1.24 \\ 0.60 & 0.53 & 0.16 & 0.32 \\ 1.50 & 1.76 & 0.08 & 0.71 \\ 0.77 & 1.48 & 0.05 & 1.23 \\ 0.74 & 1.89 & 1.70 & 0.20 \\ 0.02 & 0.83 & 0.36 & 1.71 \end{bmatrix}, \quad \theta_r(\text{in } \pi) = \begin{bmatrix} 0.25 & 1.87 & 0.73 & 0.45 \\ 1.04 & 0.73 & 0.73 & 0.16 \\ 0.39 & 0.74 & 1.51 & 1.30 \\ 0.23 & 1.52 & 1.69 & 1.40 \\ 0.77 & 0.16 & 0.55 & 1.04 \\ 1.92 & 1.63 & 1.81 & 1.71 \\ 1.80 & 1.58 & 0.07 & 1.86 \\ 1.37 & 0.44 & 0.93 & 0.16 \\ 1.85 & 1.05 & 1.19 & 0.75 \\ 0.49 & 0.64 & 1.31 & 0.35 \\ 0.20 & 1.09 & 1.28 & 0.24 \\ 0.85 & 1.39 & 0.98 & 1.20 \\ 1.13 & 0.10 & 1.13 & 1.46 \\ 0.88 & 1.92 & 1.59 & 1.08 \\ 0.84 & 1.93 & 1.68 & 1.47 \\ 1.56 & 0.25 & 1.44 & 1.96 \end{bmatrix} \quad (9.7)$$

W pierwszej symulacji porównujemy wydajność konstrukcji DB i HB. W tym celu ustawiamy $S_k = 8$ i $P_{rk} = 16, 8k$. Rysunek przedstawia sumaryczną szybkość osiągniętą przez cyfrowe i hybrydowe formowanie wiązek.

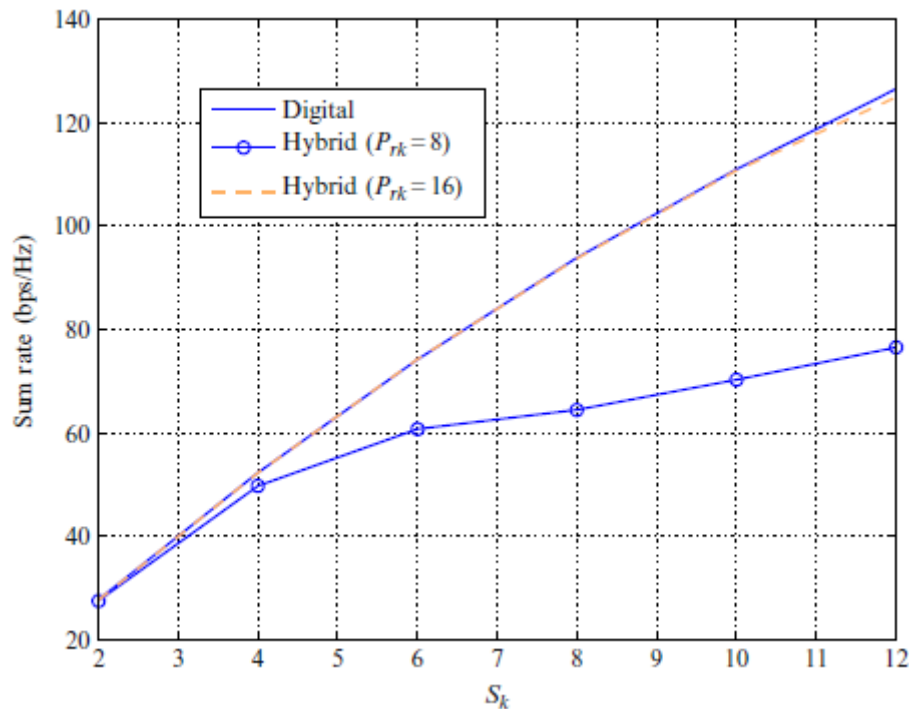


Jak widać na tym rysunku, wydajność HB jest prawie taki sam jak cyfrowy, od niskich do umiarkowanych regionów SNR. W regionach o wysokim SNR obserwuje się niewielką różnicę w wydajności. Następnie badamy wpływ P_{rk} na wydajność projektu HB. Rysunek kolejny ilustruje działanie konstrukcji HB dla różnych ustawień P_{rk} przy $\text{SNR} = \{-4 \text{ dB}, 6 \text{ dB}\}$ i $S_k=8$.



Na podstawie tej liczby można zauważyć, że wydajność konstrukcji HB poprawia się wraz ze wzrostem liczby łańcuchów RF i przetworników ADC (tj. P_{rk}), co jest oczekiwane. Na koniec analizujemy łączny

wpływ S_k i P_{rk} na wydajność cyfrowego i hybrydowego formowania wiązki. W tym celu na rysunku poniżej przedstawiono zachowanie sygnału cyfrowego i HBs przy SNR = 4 dB.



Jak widać, osiągnięte sumaryczne szybkości cyfrowych i hybrydowych formowania wiązki rosną wraz ze wzrostem S_k . Dzieje się tak, ponieważ wraz ze wzrostem S_k wzrasta również $P_{\max} = KS_k$. Również różnica w wydajności między hybrydowym i cyfrowym formowaniem wiązki maleje wraz ze spadkiem S_k . Stąd, dla ograniczonej liczby łańcuchów RF i przetworników ADC (tj. P_{rk}), luka w wydajności między hybrydowym i cyfrowym kształtowaniem wiązki może zostać zmniejszona poprzez zmniejszenie liczby multipleksowanych symboli (S_k). Na podstawie wszystkich tych danych można zauważyć, że dla danej liczby łańcuchów RF i przetworników ADC różnica w wydajności między cyfrowymi i hybrydowymi formami wiązki może zostać zmniejszona poprzez zmniejszenie liczby symboli multipleksowanych. Ponadto, dla danej liczby multipleksowanych symboli, zwiększenie liczby łańcuchów RF i przetworników ADC zwiększa osiągalną sumaryczną szybkość HB, co jest oczekiwane.

MULTIPLEKSOWANIE PRZESTRZENNE

Jak stwierdzono we wstępie, pasma niskiej częstotliwości są mniej podatne na blokowanie. Ponadto mogą one łatwo ulegać dyfrakcji w wielu fizycznych materiałach i ciałach, takich jak cegły, budynki, szkło, ludzie itp. Takie zachowanie oznacza, że odbierane sygnały pasm częstotliwości mikrofalowych zawierają kilka sygnałów odbitych od różnych rozpraszaczy, które często są niezależne. Z tego powodu kanał między dowolnymi parami anten nadawczo-odbiorczych często jest niezależny w środowisku o dużym rozproszeniu. W takim środowisku pojemność kanału bezprzewodowego MIMO można wyrazić jako:

$$C_{\text{MIMO}} = \alpha \min(M, N) \log_2(1 + \gamma(M, N)) \quad (8)$$

gdzie N (M) to liczba anten nadawczych (odbiorczych), $\gamma(M, N)$ to SINR osiągnięty przez łącze MIMO, który zależy od liczby anten, a α zależy od schematu kodowania transmisji. Jak widać na podstawie tego wyrażenia, C_{MIMO} rośnie liniowo z $\min(M, N)$. Zheng i Tse pokazali, że różne schematy kodowania

dają różne wartości α i $\gamma(M, N)$. W szczególności dowolne kodowanie schematu zwiększający jeden zmniejsza drugi i na odwrót. Teraz, jeśli weźmiemy pod uwagę transmisję niekodowaną, na której skupiamy się w tym rozdziale, zysk uzyskany przez transmisję $\min(M, N)$ symboli określa wzmocnienie multipleksowania, podczas gdy $\gamma(M, N)$ określa wzmocnienie różnorodności (tj. niezawodność łącza) kanału MIMO. Na podstawie C_{MIMO} można zaobserwować, że gdy medium transmisyjne ma akceptowalną niezawodność łącza $\gamma(M, N)$, korzystne jest wykorzystanie wszystkich dostępnych stopni swobody do transmisji $\min(M, N)$ symboli na okres symbolu, aby uzyskać lepszą C_{MIMO} . W powyższych sekcjach omówiliśmy, że systemy mmWave są odpowiednie do rozmieszczenia wielu anten (rzędu 100–1000). Ponadto działają w bardzo wysokich pasmach częstotliwości, w których dostępne szerokości pasma są znacznie większe. Potencjalne pasma widma mmWave to: 28 GHz (27,5–29,5 GHz), 38 GHz (36–40 GHz), 60 GHz (57–64 GHz), pasmo E (71–76 i 81–86 GHz) oraz W pasma (92–95 GHz). Motywuje to badaczy do rozważenia transmisji przestrzennych multipleksowanych dla kanałów mmWave w celu znacznego zwiększenia przepustowości systemu, ponieważ N (M) jest niezwykle duży. Jednak oczekuje się, że kanały mmWave będą lustrzane i mają niską rangę, a ich charakterystyka statystyczna może nie pomóc im w wykorzystaniu wszystkich dostępnych przestrzennych stopni swobody. Z tego powodu systemy mmWave mają trudności z transmitowaniem wielu strumieni w ten sam zasób czasowo-częstotliwościowy, co w konsekwencji ogranicza zyski z multiplikacji. Multipleksowanie przestrzenne zapewnia wzmocnienie multipleksowania, które zwiększa przepustowość transmisji poprzez podzielenie wychodzącego strumienia sygnału na wiele części, gdzie każdy element jest transmitowany jednocześnie i równolegle na tym samym kanale przez różne anteny. Ponadto, aby przestrzenna multipleksowana transmisja działała dobrze, kanał musi zapewniać wystarczającą dekorelację między różnymi blisko rozmieszczonymi antenami, co może nie mieć zastosowania w większości praktycznych systemów, takich jak mmWave w pomieszczeniach. Jednym ze sposobów wykorzystania transmisji multipleksowanej przestrzennie jest umożliwienie aplikacji wieloużytkownikowych systemu mmWave, w których nadajnik przesyła różne strumienie danych do różnych UE, które nie pokrywają się geograficznie. Jednak takiego podejścia nie można zastosować, gdy dwa UE są rozmieszczone bliżej siebie. W wielu innych pracach próbowano wykorzystać multipleksowanie przestrzenne dla kanałów mmWave dla systemu pojedynczego użytkownika. Torkildson omówił możliwość włączenia multipleksowania przestrzennego dla wewnętrznych kanałów mmWave poprzez modelowanie kanałów wewnętrznych jako trzech promieni; bezpośrednia linia wzroku i fale odbite z lewej i prawej strony. Za pomocą symulacji numerycznej potwierdzono, że zdolność multipleksowania takiego systemu zależy od względnych pozycji węzła nadajnika i odbiornika. Ponadto, nawet jeśli mamy tylko kanał LOS, liczbę dominujących wartości osobliwych kanału powyżej danego progu można zmaksymalizować poprzez odpowiednie dobranie pozycji anten węzłów nadawczych i odbiorczych, co w konsekwencji pomaga zwiększyć przepustowość kanału. W przypadku kanału LOS liczba kanałów multipleksowanych przestrzennie zależy od struktury szyku antenowego. Na przykład maksymalna liczba kanałów multipleksowanych przestrzennie zmienia się jako odwrotność λ i λ^2 odpowiednio dla ULA i prostokątnych układów antenowych. Aby umożliwić przestrzenne multipleksowanie LOS, wykorzystano zasady optyki o ograniczonej dyfrakcji. A dla danego zestawu anten o rozmiarze $N \times N$ ($N \times 1$) i zasięgu łącza R jest to minimalna odległość między dwiema antenami konieczna może być wybranie elementów jako:

$$D = \sqrt{\frac{R\lambda}{N}} \quad (9)$$

Ta odległość zapewnia, że kątowa separacja elementów anteny nadajnika jest większa lub równa rozdzielczości kątowej układu odbiornika. Ma to takie samo znaczenie, jak zwiększenie rangi kanału LOS (tj. Zwiększenie wzmocnienia multipleksowania przestrzennego) [19]. Multipleksowanie przestrzenne LOS zostało zademonstrowane wu Sheldona poprzez zastosowanie macierzy o rozmiarze 4 przy częstotliwości nośnej 60 GHz.

SZACOWANIE KANAŁU

Kształtowanie wiązki i multipleksowanie przestrzenne są kluczowymi technikami umożliwiającymi skalowanie przepustowości łączy systemów mmWave. Powszechnie wiadomo, że wydajność kształtowania wiązki i multipleksowania przestrzennego w dużym stopniu zależy od dostępności CSI, a lepszą wydajność uzyskuje się, gdy akwizycja CSI jest wykonywana niezawodnie. Ogólnie rzecz biorąc, estymacja kanału może być przeprowadzana przy użyciu podejścia z dupleksem z podziałem częstotliwości (FDD) lub dupleksem z podziałem czasu (TDD). W podejściu FDD najpierw nadajnik wysyła sekwencje pilota do odbiornika. Następnie odbiornik szacuje swój własny kanał. Wreszcie oszacowany kanał jest przesyłany zwrótnie do nadajnika za pośrednictwem kanału sprzężenia zwrotnego. W systemie TDD najpierw odbiornik przesyła własne symbole pilota do nadajnika. Następnie nadajnik oszacuje kanał i przypisze koniugat szacowanego kanału jako kanał między nadajnikiem a odbiornikiem. System TDD jest stosowany, gdy CSI między nadajnikiem a odbiornikiem ma charakterystykę wzajemności ze swoim kanałem zwrotnym, co jest widoczne, gdy oba łącza wykorzystują to samo pasmo częstotliwości. Jednak w systemie FDD kanały łącza w dół i łącza w górę generalnie wykorzystują wystarczająco oddzielone pasma częstotliwości. Istnieje wiele podejść do pozyskiwania danych CSI opartych na TDD i FDD dla systemów MIMO. Spośród tych metod, najmniejszy kwadrat (LS) i minimalny błąd średniokwadratowy (MMSE) są powszechnie stosowane, ponieważ są ogólne i proste i zostały omówione w następnej sekcji. Aby rozwinąć metody estymacji kanałów LS i MMSE, rozważmy oszacowanie kanału oparte na TDD dla systemu wielodostępnego, w którym stacja BS wyposażona w anteny N obsługuje K zdecentralizowanej pojedynczej anteny UE, a CSI jest szacowane w BS. W przypadku tego systemu można zastosować K okresów treningowych do oszacowania CSI wszystkich UE. Odebrany sygnał w j -tym czasie trwania szkolenia podaje:

$$y_{ij} = \sum_{i=1}^K h_i p_{ij}^* + n_{ij} \quad (10)$$

gdzie K to liczba UE, p_{ij}^* to symbol pilota i -tego UE w czasie symbolu j , $n_{ij} \in \mathbb{C}^{N \times 1}$ to odebrany wektor szumu w czasie symbolu j , a $h_i \in \mathbb{C}^{N \times 1}$ to kanał między i -tym UE i BS, który można zamodelować jako:

$$\mathbf{h}_k = \mathbf{R}_k^{1/2} \tilde{\mathbf{h}}_k$$

gdzie wpisy $\mathbf{h}_k = \mathbf{R}_k^{1/2} \tilde{\mathbf{h}}_k$ są modelowane jako niezależne i identycznie rozłożone (i.i.d.) zerowe średnie cyklicznie symetryczne zespolone zmienne losowe Gaussa (ZMCSCG), każda z jednostkową wariancją, a $\mathbf{R}_k \in \mathbb{C}^{N \times N}$ jest dodatnią półokreśloną macierzą kowariancji kanału. Wydając symbole treningowe K , stacja BS otrzymuje następujący sygnał:

$$\mathbf{Y}_t = \mathbf{H}\mathbf{P}^H + \mathbf{N}_t$$

gdzie $\mathbf{H} = [h_1, h_2, \dots, h_k]$, $\mathbf{Y}_t = [y_{t1}, y_{t2}, \dots, y_{tk}]$, $\mathbf{P} = [p_1, p_2, \dots, p_k]$, $\mathbf{N}_t = [n_{t1}, n_{t2}, \dots, n_{tk}]$ jest szum addytywny modelowany jako $\mathbf{N}_t \sim \mathcal{CN}(\mathbf{0}, \sigma_t^2 \mathbf{I})$ i $p_i = [p_{i1}, p_{i2}, \dots, p_{ik}]^T$. Dla uproszczenia wpisy \mathbf{P} są wybierane z konstelacji spełniającej $\mathbf{P}^H \mathbf{P} = \mathbf{I}$. Na przykład w systemie LTE zastosowano sekwencje Zadoff Chu. BS może oddzielić kanały każdego UE poprzez pomnożenie całego odebranego sygnału \mathbf{Y} przez \mathbf{P} , tj.

$$\mathbf{E} \triangleq \mathbf{Y}_t \mathbf{P} = \mathbf{H} + \mathbf{N}_t \mathbf{P}, \Rightarrow \mathbf{e}_k = \mathbf{h}_k + \tilde{\mathbf{n}}_k \quad (11)$$

gdzie $\tilde{\mathbf{n}}_k = \mathbf{N}_t \mathbf{P}_k$. Jak widać z równania. (11), \mathbf{e}_k nie zawiera h_i , $i \neq k$. Podejście do estymacji kanału LS traktuje \mathbf{e}_k jako oszacowanie \mathbf{h}_k . Jednak metoda MMSE obejmuje dodatkowe kroki w celu oszacowania k-tego kanału z \mathbf{e}_k . Odbywa się to poprzez wprowadzenie macierzy MMSE \mathbf{W}_k^H dla k-tego UE i wyrażenie szacowanego kanału $\hat{\mathbf{h}}_k$ jako:

$$\hat{\mathbf{h}}_k = \mathbf{W}_k^H \mathbf{e}_k \quad (12)$$

gdzie \mathbf{W}_k jest tak zaprojektowane, że średni błąd kwadratowy (MSE) między $\hat{\mathbf{h}}_k$ i \mathbf{h}_k jest zminimalizowany w następujący sposób:

$$\begin{aligned} \xi_k &= \text{tr}\{\mathbf{E}\{|\hat{\mathbf{h}}_k - \mathbf{h}_k|^2\}\} = \text{tr}\{\mathbf{E}\{|\mathbf{W}_k^H (\mathbf{h}_k + \tilde{\mathbf{n}}_k) - \mathbf{h}_k|^2\}\}, \\ &= \text{tr}\{(\mathbf{W}_k^H - \mathbf{I}_N) \mathbf{R}_k (\mathbf{W}_k^H - \mathbf{I}_N)^H + \sigma_t^2 \mathbf{W}_k^H \mathbf{W}_k\} \end{aligned} \quad (13)$$

gdzie $\text{tr}(\cdot)$ i $\mathbf{E}\{\cdot\}$ są odpowiednio operatorami śledzenia i oczekiwania. Optymalny \mathbf{W}_k otrzymujemy z gradientu ξ_k jako:

$$\frac{\partial \xi_k}{\partial \mathbf{W}_k^H} = 0, \Rightarrow \mathbf{W}_k^* = (\mathbf{R}_k + \sigma_t^2 \mathbf{I})^{-1} \mathbf{R}_k \quad (14)$$

Jak widać z wyrażen (11) i (12), różnica między estymatorami kanałów LS i MMSE polega na tym, że ten pierwszy nie wykorzystuje informacji statystycznych z macierzy kowariancji szumu i kanału, podczas gdy drugi wymaga statystyki szumu i macierz kowariancji kanałów. Na podstawie tych estymatorów można również zauważyć, że informacji CSI wszystkich UE nie można odsprzęgnąć, jeżeli zastosujemy mniej niż K okresów treningowych, co jest równe całkowitej liczbie anten nadawczych kanału łączy w górę. Z powyższych dyskusji można również zauważyć, że przynajmniej N symboli pilota jest potrzebnych do efektywnego pozyskania CSI, jeśli szkolenie jest wysyłane z BS. W systemie komunikacji mikrofalowej, ponieważ K i N są porównywalne, wysyłanie symboli pilota do stacji BS lub UE nie spowoduje znaczącej różnicy. Ponadto dostarczanie informacji zwrotnej CSI do UE lub BS nie jest kosztowne. Z tego powodu można zastosować system TDD lub FDD. Jednak gdy $N \gg K$ (tj. Masywny system MIMO), wysyłanie pilotów z UE jest ekonomiczne, co w konsekwencji prowadzi do promowania podejścia CSI opartego na TDD. Chociaż istnieją obszerne badania dotyczące różnych metod akwizycji CSI systemu MIMO, większości metod nie można zastosować bezpośrednio w systemach mmWave. Główne powody to:

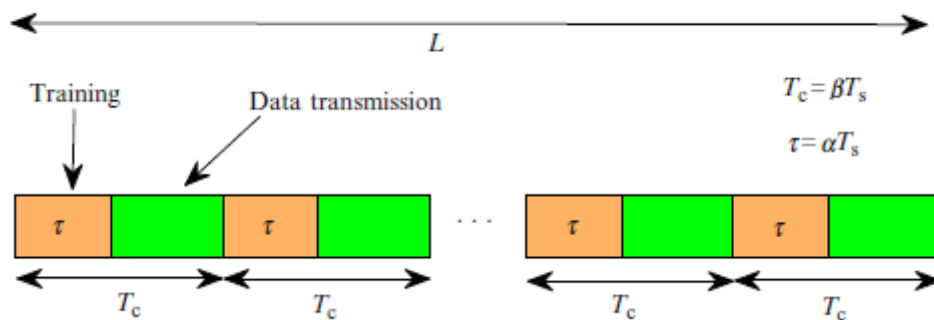
- Przesunięcie Dopplera skaluje się liniowo wraz z częstotliwością, tak więc czas koherencji kanałów mmWave jest mniej więcej o rząd wielkości mniejszy niż porównywalnych pasm mikrofalowych. Na przykład w typowych warunkach czas koherencji kanału może wynosić 500 i 35 μs , gdy system jest wdrażany na częstotliwościach nośnych, odpowiednio, 2 i 28 GHz. Z tego powodu oczekuje się, że

ogólne wyzwania związane z narzutem uczenia się CSI w kanałach mmWave będą znacznie wyższe niż w przypadku mikrofalowych masywnych systemów MIMO dla porównywalnego środowiska mobilności. Co więcej, kanały mmWave doświadczają poważnych blokad, gdzie akwizycja CSI staje się trudna, ponieważ prawdopodobieństwo wyłączenia łącza jest znacznie większe z powodu niewspółosiowości kształtowania wiązki (łącza).

- Prawie wszystkie konwencjonalne systemy MIMO zakładają, że liczba rozmieszczonych łańcuchów RF jest taka sama jak w przypadku anten, co może nie być wykonalne w zastosowaniach mmWave, ponieważ takie systemy będą wyposażone w duże tablice antenowe. Ponadto, gdy liczba łańcuchów RF w BS jest znacznie mniejsza niż liczba anten BS, nie można zastosować metody estymacji kanału ortogonalnego. Wynika to z faktu, że odbiornik straci pewien stopień swobody w rzetelnym szacowaniu współczynników CSI.

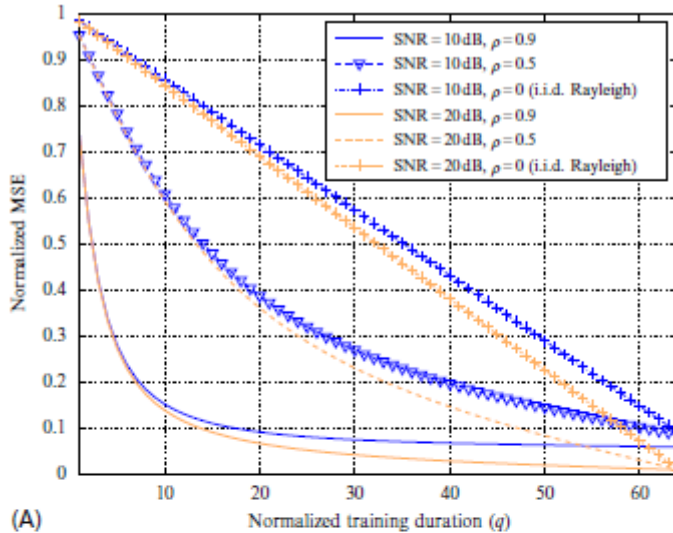
Z powyższych dwóch powodów podjęto szereg wysiłków badawczych w celu opracowania różnych technik umożliwiających wydajne oszacowanie kanału. Ogólnie rzecz biorąc, jakość estymatora kanału można poprawić, wydłużając okres szkolenia. Ponieważ jednak kanały praktyczne mają ograniczony czas koherencji, gdzie kanał jest traktowany jako prawie stały, nie można wybrać dużego okresu szkolenia. Ponadto, ponieważ uzyskany CSI jest używany do transmisji danych, dłuższy czas akwizycji CSI doprowadzi do zmniejszenia przepustowości sieci dla danego czasu koherencji. Z tego powodu akwizycja CSI może wymagać wydajnego wykonania, biorąc pod uwagę zarówno dokładność, jak i czas trwania szkolenia. Jak wspomniano, można zastosować hybrydową architekturę analogowo-cyfrową pojedynczego PS lub sparowanych PS w systemach mmWave wyposażonych w mniej łańcuchów RF niż anteny. Problem estymacji kanału dla systemów mmWave został sformułowany jako wykrywanie kompresyjne w Bajwa i in., Alkhateeb i in. oraz Malloy i Nowak [66–68], stosując pojedynczą hybrydową architekturę analogowo-cyfrową PS. Praca Alkhateeba stosuje podejście adaptacyjnej książki kodowej do projektowania swoich szkoleń dla modeli macierzy anten ULA. Ramasamy zaproponował podejście do wykrywania skompresowanego kanału z architekturą analogowego kształtowania wiązki, w którym podejście to próbuje oszacować wstępnie zdefiniowaną liczbę ścieżek przy użyciu siatki częstotliwości przestrzennych, po czym następuje dalsze dostrojenie. Podobnie, skompresowane wykrywanie kanałów z przypadkowymi sekwencjami treningowymi zostało zapewnione w architekturze hybrydowej u Lee, Bogale i inni rozważali ostatnio wspólną estymację kanału i formowanie wiązki, biorąc pod uwagę czas koherencji kanału, stosując sparowaną hybrydową architekturę analogowo-cyfrową PS. Akdeniz przeprowadził szeroko zakrojone eksperymenty plenerowe w celu scharakteryzowania kanałów mmWave. W artykule omówiono niektóre z charakterystyk kanałów mmWave w dużej skali, takie jak kąt nadejścia (AoA), kąt wyjścia (AoD), kąt wiązki rozproszonej wokół środka oraz profil mocy i opóźnienia grupowego każdego klastra, gdzie wszystko to zmienia się powoli w porównaniu z zanikaniem na małą skalę. Wynik ten potwierdza, że macierze kowariancji kanałów mmWave byłyby prawie stałe przez znaczny czas trwania symbolu. Z drugiej strony Eliazi i inni niedawno zaproponowali różne metody estymacji macierzy kowariancji kanałów dla kanałów fal milimetrowych, wykorzystując hybrydową architekturę analogowo-cyfrową. Ponadto, ponieważ systemy mmWave są wyposażone w duże tablice antenowe, podejście TDD jest najbardziej rozsądne. Te pomysły motywowały Bogalea w celu wykorzystania okresowej struktury ramek estymacji kanału opartej na TDD i transmisji danych poprzez wykorzystanie informacji o kowariancji kanałów mmWave, jak pokazano na rysunku. W przypadku tej struktury ramowej, w drugim artykule zbadano kompromis w zakresie przepustowości szkolenia dla wieloużytkownikowych systemów mmWave MIMO z ograniczoną liczbą łańcuchów RF. Praca Bogalea wykorzystuje fakt, że optymalny czas trwania szkolenia, który maksymalizuje całkowitą przepustowość sieci, zależy od działającego SNR, dostępnej liczby łańcuchów RF, czasu koherencji kanału (T_c) i macierzy kowariancji wszystkich UE. W szczególności wykazano, że problem optymalizacji czasu szkolenia można

sformułować jako problem maksymalizacji wklęsłej, w którym jego globalne optymalne rozwiązanie można skutecznie uzyskać przy użyciu dostępnych na półce narzędzi do optymalizacji wypukłych, takich jak CVX. Poniżej przedstawiamy wyniki liczbowe dla wydajności poziomego łącza estymacji kanału mmWave, która wykorzystuje sparowaną hybrydową architekturę analogowo-cyfrową PS. Uwzględniamy osiągniętą przepustowość na poziomie łącza dla stałego czasu koherencji kanału. W związku z tym rozważamy sieć bezprzewodową pojedynczego użytkownika, w której stacja BS wyposażona w anteny N obsługuje jedną pojedynczą antenę UE. Liczba łańcuchów RF w BS jest ustawiona na 1. Rozważamy strukturę ramki, jak pokazano na rysunku, gdzie T_c jest czasem koherencji kanału, τ jest czasem uczenia, a $T_c - \tau$ jest czasem trwania transmisji danych, a $\beta = 2T_0$, $\alpha = \theta$ i T_0 zostały wprowadzone dla wygody.

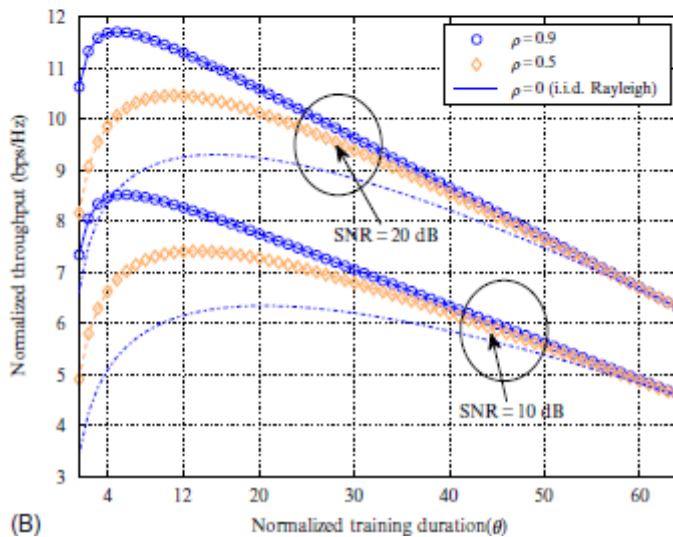


Macierz kowariancji kanału UE R jest pobierana z szeroko stosowanego modelu korelacji wykładniczej jako $R_{(i,j)} = \rho^{|i-j|}$, który zakłada się, że jest prawie stały dla wystarczająco dużej liczby ramek L , gdzie $0 \leq \rho < 1$. Wybraliśmy ten model, ponieważ oddaje charakterystykę różnych środowisk propagacji. Na przykład kanały mmWave są często reprezentowane przez macierze kowariancji kanałów niskiego rzędu, które można modelować, wybierając duże ρ .

Znormalizowana estymacja kanału MSE dla $N = 64$ i $T_0 = 64$, $\rho = \rho_1$ i różnych θ i SNR jest wykreślona na rysunku A.



(A)



(B)

Jak widać na tej figurze, zwiększenie θ poprawia jakość estymatora kanału poprzez zmniejszenie MSE, co jest oczekiwane. Z drugiej strony, dla danego operacyjnego SNR, czas trwania szkolenia wymagany do osiągnięcia danego docelowego MSE zmniejsza się wraz ze wzrostem ρ . Wynik ten jest oczekiwany, ponieważ wraz ze wzrostem ρ λ_i/λ_1 z $\lambda_1 \geq \lambda_2, \dots, \lambda_N$ maleje wraz ze wzrostem i , co w konsekwencji pomaga skrócić czas trwania treningu. Rysunek B przedstawia osiągalną przepustowość w funkcji czasu trwania treningu dla różnych SNR i ρ . Jak widać, pomimo podejścia do estymacji kanału stosowanego w systemach mikrofalowych, optymalny czas trwania szkolenia hybrydowego analogowo-cyfrowego estymatora kanału niekoniecznie jest taki sam dla wszystkich ρ . W szczególności optymalny czas trwania treningu zmniejsza się, gdy ranga macierzy kowariancji kanału maleje (tj. Wraz ze wzrostem ρ). Na przykład, gdy SNR = 10dB, optymalny czas trwania treningu wynosi około $4T_s$, $12T_s$ i $20T_s$ odpowiednio dla $\rho = 0,9$, $0,5$ i 0 (tj. i.i.d. zanikającego kanału Rayleigha). Z drugiej strony krótszy jest okres szkolenia stosowany, gdy kanał UE jest silnie skorelowany, co w konsekwencji skutkuje wyższą znormalizowaną przepustowością. Wyniki te potwierdzają, że okres szkolenia należy dobrać ostrożnie, gdy stacja bazowa ma mniej łańcuchów RF niż anteny, co często ma miejsce w przypadku systemów mmWave. Jak widać na rysunku, procedura szacowania kanału dla systemów mmWave różni się od procedury dla systemów mikrofalowych. Ponadto istnieje kompromis między czasem trwania szkolenia a osiągalną przepustowością, gdy liczba łańcuchów RF jest ograniczona i stosowana jest hybrydowa estymacja kanałów analogowo-cyfrowych i formowanie wiązki.

PROJEKTOWANIE FAL

Wydajność na poziomie łącza kanału bezprzewodowego jest również określana przez sposób projektowania przebiegów przesyłanych sygnałów. W rzeczywistości różne projekty przebiegów prowadzą do różnych osiągnięć systemu pod względem estymacji kanału i kształtowania wiązki, co z kolei daje różne możliwości poziomu łącza. Pod tym względem wydajność dowolnego kształtu fali zależy od jego zdolności do wydajnej realizacji sygnalizacji i multipleksowania. W istniejących standardach bezprzewodowych istnieje wiele podejść do projektowania fal: wielodostęp z podziałem czasu (TDMA) i wielodostęp z podziałem częstotliwości (FDMA) zostały zastosowane w systemach 2G ze względu na ich zdolność do multipleksowania symboli informacyjnych. Widmo rozproszone (CDMA) staje się bardziej dominujące dla systemu 3G. OFDM został wybrany przez szereg najnowszych standardów bezprzewodowych, ponieważ ma kilka zalet w porównaniu do swoich odpowiedników TDMA, FDMA i CDMA. Wszystkie te techniki wykorzystują podejście transmisji Nyquista, w którym tylko jeden element zawartości informacji jest włączony na czas trwania symbolu. Ponieważ przyszła sieć 5G wymaga znacznej przepustowości, najnowsza literatura próbuje zasugerować strategię transmisji, która może obsłużyć więcej treści informacyjnych na okres symbolu. Pod tym względem rozważana jest transmisja szybsza niż Nyquist (FTN). Pierwsza praca nad transmisją FTN została zaproponowana przez Mazo w 1975 r., w której twierdzi się, że możliwe jest zapewnienie prędkości transmisji wykraczającej poza to, co jest zdefiniowane przez kryterium Nyquista, bez narzucania dodatkowej przepustowości. Wielkość prawdopodobieństwa błędu przesyłanego symbolu odgrywa fundamentalną rolę w zrozumieniu kanału komunikacyjnego, który może być bezpośrednio powiązany z minimalną odległością euklidesową d_{\min} między dwoma różnymi symbolami. Kluczową ideą Mazo jest to, że dla scenariusza z czasem trwania symbolu Nyquista T_{Ny} i filtrem kształtującym impuls sinc, d_{\min}

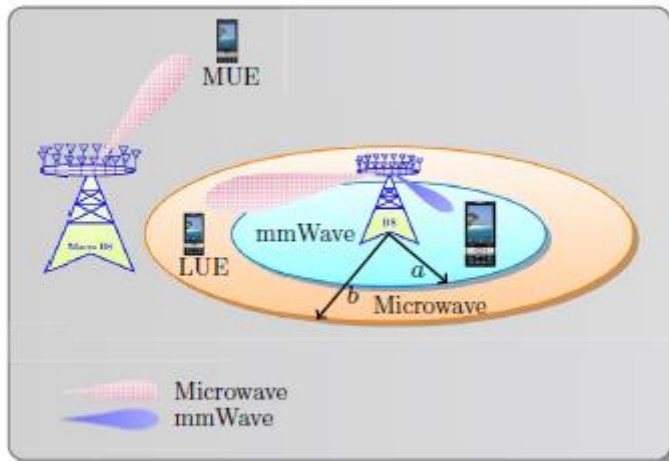
+ będzie takie samo, gdy czas trwania symbolu jest ustawiony na τT_{Ny} z $\tau \in [0,802,1]$. Wynik ten uzasadnia, że można nadal zachować takie samo prawdopodobieństwo błędu, jak w przypadku transmisji Nyquista, wybierając $\tau \approx 0,82$, co z kolei pomaga poprawić efektywność przepustowości o $1/0: 802 \approx 25\%$. Wykazano, że różne filtry kształtujące impuls mają różne granice Mazo (tj. minimalne τ). Na przykład minimalne τ z filtrem kosinusowym ze zwiększoną pierwiastkiem i nadmierną szerokością pasma $\beta = 0,3$ daje 42% wzrost wydajności widmowej (tj. $\tau_{\min} \approx 0,703$). Chociaż transmisja FTN osiąga lepszą przepustowość kanału niż transmisja Nyquista, wymaga nieliniowej operacji dekodowania, co w praktyce nie jest pożądane. Z tego powodu transmisja Nyquist z obsługą OFDM jest nadal popularna w kilku standardach bezprzewodowych LTE, LTE-A, WiFi, WiMax, DVB2 itd. Główne zalety przebiegów OFDM to:

1. Aby osiągnąć bardzo dużą szybkość transmisji danych, wymagana jest duża szerokość pasma, które z kolei zapewnia kanał selektywny częstotliwościowo. Podejście projektowe OFDM pomaga przekształcić kanały selektywne częstotliwościowo na równoległy, płaski kanał zanikający. W konsekwencji pomoże to niezależnie zbadać składowe sygnały, w których występują płaskie zanikające kanały (tj. Podnośne).
2. Kształt fali OFDM można generować za pomocą operacji odwrotnej szybkiej transformaty Fouriera (IFFT), a przesyłany sygnał można odzyskać za pomocą operacji szybkiej transformacji Fouriera (FFT). Ponieważ operacje FFT i IFFT mogą być realizowane efektywnie, OFDM jest atrakcyjny z punktu widzenia implementacji.
3. Ponieważ OFDM umożliwia oddzielną obsługę każdej podnośnej, różne podnośne mogą być przydzielane różnym użytkownikom, co w konsekwencji umożliwia multipleksowanie o niskiej złożoności. Ponadto można ją rozszerzyć na MIMO OFDM dla wielu użytkowników na podstawie podnośnej.

Transmisja OFDM ma jednak jedną zasadniczą wadę. Ponieważ nadajnik wykorzystuje operację IFFT na transmitowanych symbolach (tj. wstępne kodowanie), taka operacja znacznie zwiększy PAPP transmitowanego kształtu fali, zwłaszcza gdy rozmiar IFFT jest duży. W konsekwencji będzie to wymagało wysokiej jakości PA, które jest często drogie, zwłaszcza w przypadku terminali o małej mocy transmisji. Ponadto, jeśli używa się taniego PA, niektóre przesyłane sygnały zostaną obcięte, co z kolei spowoduje utratę ortogonalności każdej podnośnej. Ponadto duży problem z PAPP skutkuje poważnym pogorszeniem wydajności BER. Aby to obejść, zaproponowano różne zmodyfikowane transmisje OFDM, z których transmisja z pojedynczą nośną jest najczęściej stosowanym podejściem (np. W transmisji w kanale up-link). Spowoduje to powstanie pytania, czy OFDM nadal można wybrać jako realne rozwiązanie w zastosowaniach mmWave. Jak stwierdzono w powyższych sekcjach, znaczna szerokość pasma jest dostępna w pasmach częstotliwości mmWave (np. Szerokość pasma około 12,9 GHz jest dostępna w paśmie E), co oznacza, że można uzyskać komunikację o wysokiej przepustowości w falach mm. Z drugiej strony, ponieważ pasma częstotliwości mmWave mają bardzo małe czasy koherencji kanału, czas trwania szczeliny każdego bloku transmisji będzie bardzo niski w porównaniu z pasmami częstotliwości mikrofalowych. Na przykład czas trwania szczeliny w obecnym systemie LTE wynosi około 0,5 ms, który jest w stanie obsłużyć 6–7 symboli OFDM, z których każdy ma czas trwania 66,7 μ s. Co więcej, PA o częstotliwości mikrofalowej są tańsze niż PA do aplikacji mmWave. Ponieważ pasma mmWave wykorzystują masywne tablice antenowe, wdrożenie wysokiej jakości PA do aplikacji mmWave sprawi, że transmisja OFDM nie będzie preferowana. Ponadto, ponieważ czas trwania szczeliny jest bardzo mały, a szerokość pasma jest wysoka, prosty schemat TDMA (FDMA) mógłby być praktycznie wykonalny dla systemów z wieloma użytkownikami, przy jednoczesnym włączeniu systemu z pojedynczą nośną, aby umożliwić projektowanie odbiornika o małej złożoności. Wszystkie te powody przemawiają za zastosowaniem podejść do projektowania przebiegów TDMA (OFDM) z pojedynczą nośną dla systemów mmWave.

STRATEGIA DOSTĘPU

Wydajność na poziomie łącza systemów mmWave zależy również od strategii dostępu zastosowanej w sieci. W związku z tym naukowcy zaproponowali łączność podwójną falą, aby osiągnąć korzyści zarówno z pasm mikrofalowych, jak i pasm mmWave dla nałożonych małych komórek mmWave (tj. Mała komórka jest pokryta makro BS). Główną ideą łączności dwufalowej jest to, że każde UE działa jednocześnie w pasmach częstotliwości mikrofalowej i mmWave. W ten sposób płaszczyzny sterowania i danych mogą zostać rozdzielone, tak aby krytyczne dane sterujące były przesyłane przez niezawodne łącza mikrofalowe między UE i makro BS, podczas gdy szybka komunikacja danych między UE a małokomórkowymi stacjami BS jest realizowana w pasmach częstotliwości mmWave. W konsekwencji można utrzymać niezawodną i stabilną łączność, czerpiąc korzyści z pasm mmWave. Na przykład dwuzakresowa mała komórka z priorytetową i warstwową strategią asocjacji komórek, jak pokazano na rysunku, może wykorzystać zalety tych różnych pasm częstotliwości.



W tym projekcie obszar pokrycia jest podzielony na trzy regiony, w których urządzenia UE w obszarze wewnętrznym (tj. promień a) / obszarze środkowym (tj. promień b) są obsługiwane przez małą komórkę przy częstotliwości fali mm / mikrofal y. Urządzenia UE w obszarze zewnętrznym (tj. O promieniu $> b$) są obsługiwane przez makrokomórkę w pasmach częstotliwości mikrofalowych. Tutaj region zewnętrzny działa w pasmach częstotliwości mikrofalowych, ponieważ te pasma umożliwiają komunikację na duże odległości w celu wspierania mobilności. Parametry a i b można wybrać lub zoptymalizować adaptacyjnie na podstawie różnych kryteriów. Ta dwuzakresowa konstrukcja małokomórkowa izoluje urządzenia UE o falach mm i częstotliwości mikrofalowej (tj. UE o falach mm w promieniu a nie doświadczają żadnych zakłóceń ze strony makro BS, co jest pożądane w celu zwiększenia przepustowości na poziomie łącza). Akdeniz przeanalizował pojemność układów pikokomórkowych mmWave. Wykazano, że sieci mmW będą działać w warunkach skrajnie ograniczonej mocy (szumu), w których pełne przestrzenne i szerokopasmowe stopnie swobody nie są w pełni wykorzystywane. Z tego powodu wzorzec kształtowania wiązki jednego UE nie będzie silnie kolidował z wzorcem innych UE . Spowodowało to wykorzystanie częstotliwości mmWave zarówno w sieciach dostępowych, jak i dosyłowych. W ten sposób można osiągnąć większą elastyczność wdrażania [26]. W tym zakresie przeprowadzono szereg prac badawczych. Taori i Sridharan rozważali rozwiązanie bezprzewodowe typu backhaul w paśmie, w którym backhaul i dostęp są multipleksowane w tym samym paśmie częstotliwości poprzez harmonogramowanie w oparciu o TDD. Problemy alokacji zasobów dla wspólnego projektowania sieci dostępowych i dosyłowych zostały zbadane u Bernardosa. Wyzwania projektowe i wdrożeniowe związane ze wspólnym dostępem do mmWave i komunikacją typu backhaul zostały również omówione i Dehosa. Wszystkie te czynniki pomagają poprawić wydajność na poziomie łącza mmWave.

WNIOSKI

Omówiono kluczowe techniki wpływające na wydajność poziomu łącza przyszłych systemów bezprzewodowych 5G opartych na mmWave. W szczególności szeroko omówiono formowanie wiązki, multipleksowanie przestrzenne, estymację kanałów, projektowanie fal i strategię dostępu. Wykazano, że przepustowość na poziomie łącza systemów mmWave zależy nie tylko od liczby rozmieszczonych anten, ale także od liczby łańcuchów RF. W rzeczywistości najlepszą wydajność uzyskuje się, gdy systemy mmWave są zaprojektowane w taki sposób, że liczba łańcuchów RF jest taka sama, jak w przypadku anten. Jednak taka strategia projektowa będzie bardzo kosztowna, ponieważ systemy mmWave są prawdopodobnie wyposażone w duże macierze antenowe. Ogólnie rzecz biorąc, koszt systemu mmWave można obniżyć do akceptowalnego poziomu, wykorzystując hybrydową architekturę analogowo-cyfrową kosztem utraty wydajności. W powyższych sekcjach pokazano, że

wdrożenie hybrydowej architektury analogowo-cyfrowej znacząco zmienia kształtowanie wiązki, multipleksowanie przestrzenne, estymację kanału i przebieg transmisji. Omówiono również pokrótce, że komunikacja przestrzenna multipleksowana jest możliwa dla kanałów LOS mmWave, co ułatwia komunikację w pomieszczeniach o bardzo dużej szybkości transmisji danych. Z drugiej strony, zastosowanie architektury, która umożliwia łączność dwufalową, jest również unikalnym aspektem komunikacji mmWave, która zapewnia różne poziomy przepustowości łącza.